

Thesis Proposal

“Measurement of the Top Quark mass in lepton+jets events with lifetime tagging”

Robert Harrington

November 16, 2004

1 Introduction

The mass of the top quark is of interest not only because the top quark is a fundamental particle in the Standard Model of particle physics, but also because the mass of the t quark, along with that of the W boson, can be used to constrain the mass of the yet-to-be-observed Higgs boson.

Proton-antiproton ($p\bar{p}$) collisions produce top quarks in top-antitop ($t\bar{t}$) pairs at the Tevatron (at $\sqrt{s} = 1.96$ TeV in Run II). Nearly 100% of these top quarks decay to a W boson and a bottom (b) quark, so the final state resulting from the decay of the top quark pair consists of jets originating from the two b quarks, and the decay products from each of the two W bosons.

The W boson has two decay modes, decaying either to two quarks (branching fraction = $2/3$), or alternately to a lepton (l) and its corresponding lepton neutrino (ν_l) ($1/9$ for each lepton flavor). Thus $t\bar{t}$ decays 45.4% of the time entirely to jets, 11.2% to final states with two leptons and jets, and 44.4% to states with single leptons and jets.

The single lepton channel is used for the top mass measurement described in this proposal. Even though statistics are higher in the channel in which both W bosons decay to jets, the measurement would have higher systematic uncertainty due to difficulties in converting jet energies to final state parton energies. The dilepton channel has 2 b -quarks in its final state and no partons from leading order interactions, but this channel has poor statistics. The compromise is to use the single lepton channel, which has 4 final state partons from leading order interactions and relatively good statistics.

The goal of this analysis is to measure the top quark mass with the most precision of any single channel analysis. To do so, we will increase the purity of the lepton+jets data sample through the use of lifetime tagging, which attempts to identify events with b -quarks, and sophisticated kinematic fitting techniques.

2 Description of detector

The measurement is performed by using data acquired by the DØ detector during Run II of the Fermilab Tevatron collider. The signature for the single lepton channel is a lepton with high transverse momentum (p_T) and at least 4 jets. Detecting this event and performing the analysis to discriminate between this event and background, both instrumental and physics, requires an understanding of every part of the detector and of the triggering system.

The tracker provides b -tagging information as well as very precise knowledge of p_T for all particles. The hadronic and electromagnetic calorimeters provide information about jets, electrons, and other electromagnetic showers. The calorimeters also allow for the reconstruction of missing transverse energy, which is assumed to be due primarily to neutrinos. The muon system provides precise identification of muons which can be matched to tracks in the central tracker and/or calorimeter. Muon scintillation counters provide the ability to trigger specifically on events containing at least one muon, which account for 1/3 of the single lepton channel events.

Muon detection has been less complete in the bottom of the DØ detector where space is limited due to structural supports. I have worked to improve this coverage by installing new muon scintillation detectors in this region [1]. Since top events occur more frequently in the central region of the detector, due to the high transverse momentum of its decay products, this improved coverage immediately impacts the number of events available for calculating the top mass. These counters still require final calibration to optimize their contribution to the overall triggering of the DØ detector, so I plan to complete this and study its impact on the top mass and other analyses.

In addition to the calibration of these specific counters, I will continue studies to improve the performance of the entire central muon scintillation system. This involves looking directly at the data being recorded by the muon system prior to combination of the raw muon time information with information from other detector systems [2]. This is necessary because information needed for the calibration is not retained at later levels of reconstruction. Calibration of these times results in greater ability to discriminate between muons coming from actual $p\bar{p}$ collisions in the detector, which will have times closely matched to the collision times, and cosmic muons with random time distributions. Improvements in these times have made it possible to have triggers which prevent events with large times from being recorded, thus improving overall efficiencies.

3 Analysis

Once events are recorded, we need to separate top events from a relatively large number of background events. The most significant background comes from the production of W bosons together with jets. In addition, instrumental backgrounds arise from QCD processes with at least 4 jets with π^0 's and γ 's misidentified as electrons, or semi-leptonic heavy flavor decays producing muons which appear to be isolated from jets.

After applying selection criteria to obtain events which most look like $t\bar{t}$ events, and then using techniques to further discriminate signal from background, the signal-to-background ratio is roughly 1/3 for the lepton+jets channel. The method used to measure the mass of the top quark must therefore be able to take into account the large background. The background is considerably reduced if one requires at least one of the jets to be identified as originating from a b -quark.

After initial preselection, we build a likelihood function which relates the probability density for observed events to the probability density for signal ($t\bar{t}$) and background (W +jets and QCD) events. This likelihood is a function of a measured observable, in this case the top mass, and it is minimized to give the most likely value of the top mass and the uncertainty in its measurement.

The Matrix Element method, used recently to recalculate the top mass using data from the previous run of the Tevatron (Run I), is based on calculating the differential cross-sections according to Fermi's Golden Rule for each event to provide discrimination between signal and background [7]. This is different from the earliest methods of Run I, which utilized differences in topological variables such as aplanarity and jet clustering to distinguish signal from background [6]. It is assumed in the Matrix Element method that either $t\bar{t}$ or W +jet production is responsible for the event. Differential cross-sections are appropriately normalized and used to determine signal and background probability densities. Parton energies are calculated using "transfer functions", which relate measured jet energies to parton energies and are determined using Monte Carlo simulation.

In order to test the validity of the method, we must calibrate it, understand the backgrounds, and understand all systematic effects associated with it. We must also understand the issues unique to using b -tagging information to improve the analysis. In order to do this, I plan to address the following issues:

1) We use Leading Order Matrix Element calculations, but in reality there are extra jets arising from gluon radiation and other second-order effects. The Monte Carlo simulation uses the ALPGEN[8] generator for generating final-state partons and PYTHIA[9] for creation of showers from those partons. In order to minimize the impact of higher order corrections in the measurement which only uses a Leading Order approximation, we: a) require exactly 4 jets in the event, and b) apply cuts, when using Monte Carlo simulated events, to ensure the jets all came from final state partons and not from gluon radiation (jet-parton matching). While the first step is trivial, the second requires further study. In order to estimate systematic uncertainties, higher-order Matrix Element calculations will be considered, as well as the effect of adding to the

measurement events with more than 4 jets.

2) Identifying jets as b -jets (b -tagging) is expected to help in signal-background discrimination. In Run I, at DØ b -tagging was possible only through the identification of semi-leptonic decays of B hadrons, but upgrades of the DØ detector for Run II have included a Silicon Microstrip Tracker (SMT) and a Central Fiber Tracker (CFT) immersed in a 2 Tesla solenoidal magnetic field. These detector components allow for very good track reconstruction close to the primary vertex (PV) of interaction. They also allow for reconstruction of secondary vertices a few millimeters away from the PV, which is a strong indication of a B hadron decay (i.e. lifetime b -tagging). The probability for a W +jets event where all 4 jets are light jets to be b -tagged is quite low ($\sim 1\%$). The probabilities for W +heavy jets events to be b -tagged is significantly higher ($\sim 50\text{-}70\%$, depending on the flavor). Therefore it is necessary to understand and correctly model all W +heavy jets backgrounds.

3) It is essential that we can predict how each sample ($t\bar{t}$, W +light jet, W +heavy jet, etc.) will contribute to a b -tagged sample. Besides the likelihood fits used to determine the top mass, likelihood fits using kinematical variables are used to determine the approximate numbers of signal, W +jets, and QCD events to include in the ensembles (described below) used in the calibration of the method. It is not yet understood how heavy-flavored events will skew the kinematic distributions, which may in turn affect this kinematic fit. This is a potential source of bias in the measurement which must be properly understood.

4) Tagging information can be used in two different ways. It can be used to assign a higher weight to those events which are likely to contain a b -tagged jet, up to and including completely eliminating those data events which do not contain a b -tagged jet. If this is done, the signal and background probabilities are calculated as it was done for the Run I measurement with no specific knowledge of the heavy content of jets in the event.

A second approach is to incorporate the tagging information into the calculation of the signal and background probabilities. As mentioned earlier, the differential cross-section is calculated according to Fermi's Golden Rule. This method determines what the quark energy would have been prior to the quark creating a jet shower. In the Run I measurement, since it wasn't known which jets came from heavy quarks, all possible combinations of light and heavy quarks creating 4 jets were used with equal weights to determine the overall signal probability. Using b -tagging information will allow the weighting of all 24 probabilities, giving more weight to the combinations that more closely match the event, based on tagging information.

In either approach, one can either keep all events in the data sample or reject events which do not have at least one b -tag.

In order to calibrate the method, ensembles are created which contain the approximate number of signal and background events after the initial pre-selection. QCD events are modeled from data, since Monte Carlo does not adequately describe QCD processes. The other processes (W +jets and $t\bar{t}$) are simulated using Monte Carlo. These ensembles are used to study systematic biases and uncertainties expected under the various scenarios being considered. In the end, the

method which results in the lowest combination of systematic and statistical error will be used for the measurement of the top mass using the DØ data.

4 Studies of Systematic Errors

Major sources of systematic uncertainty in the measurement are 1) uncertainties in the determination of the jet energies, and 2) modeling of higher order corrections (gluon radiation) in both the signal and the background processes. We plan to address the first by looking at samples of γ +jet events with at least one b -tagged jet in the event to study the difference between tagged and untagged events, and to assign an uncertainty to the tagged jet energy calibration. For the second, I will run the analysis with and without jet-parton matching applied to the Monte Carlo in order to estimate the effects on the final top mass and uncertainty. I will also look at differences between 4- and 5- jet events to quantify the effect of including jets which came from gluon radiation.

5 Proposed Thesis Committee

George Alverson
Emanuela Barberis (advisor)
Pran Nath

References

- [1] R. Harrington *et al*, DØ Note 4601, (2004).
- [2] R. Harrington *et al*, DØ Note 4435, (2004).
- [3] CDF Collaboration, B. Abbott *et al*, Phys. Rev. Lett. **74**, 2626 (1995).
- [4] DØ Collaboration, F. Abe *et al*, Phys. Rev. Lett. **74**, 2632 (1995).
- [5] DØ Collaboration, S. Abachi *et al*, Phys. Rev. D **52**, 4877 (1995).
- [6] DØ Collaboration, S. Abachi *et al*, Phys. Rev. Lett. **79**, 1197 (1997).
- [7] DØ Collaboration, V. M. Abazov *et al*, Nature 429, 638-642 (2004).
- [8] ALPGEN, M.L. Mangano *et al*, JHEP 0307:001, hep-ph/0206293 (2003).
- [9] PYTHIA 6.2, T. Sjöstrand *et al*, LU TP 01:21, hep-ph/0108264 (2003).